UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/820,671	04/08/2004	Andrew Zisserman	13058N/040714	8040

32885   7590   08/06/2008

STITES & HARBISON PLLC
401 COMMERCE STREET
SUITE 800
NASHVILLE, TN 37219

EXAMINER
STREGE, JOHN B

ART UNIT	PAPER NUMBER
2624	

MAIL DATE	DELIVERY MODE
08/06/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

	Application No.	Applicant(s)
Office Action Summary	10/820,671	ZISSERMAN ET AL.
	Examiner	**Art Unit**	
	JOHN B. STREGE	2624	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on *02 June 2008*.
2a) ☒ This action is **FINAL**.   2b) ☐ This action is non-final.
3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *1-21* is/are pending in the application.
    4a) Of the above claim(s) _____ is/are withdrawn from consideration.
5) ☐ Claim(s) _____ is/are allowed.
6) ☒ Claim(s) *1-21* is/are rejected.
7) ☐ Claim(s) _____ is/are objected to.
8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.
10) ☒ The drawing(s) filed on *12 October 2004* is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
    a) ☐ All   b) ☐ Some *   c) ☐ None of:
      1. ☐ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____.
      3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *4/01/08*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____ .

Response to Amendment

1.   The amendment received 6/02/08 has been entered in full.

Response to Arguments

2.   The Applicant's arguments regarding the objection to the drawings are persuasive, thus this objection has been withdrawn. Applicant's arguments regarding the art rejections have been fully considered but they are not persuasive.

Regarding claim 1, the Applicant argues that Jain fails to disclose the use of vectors, with each vector having a descriptor representing a co-variant region of the object. The Examiner respectfully disagrees. As discussed in the previous office action Jain discloses a feature vector describing a visual object (col. 10 line 16). Jain goes on to disclose that a visual feature is any property of an image that can be computed using computer vision or image-processing techniques, including hue, saturation, texture measures, periodicity, and orientation which can be computed over a small region in the image. Covariance is defined as varying in accordance with a fixed mathematical relationship, thus texture measures, and periodicities to name a few are covariant features. Thus col. 10 lines 5-26 reads on defining one or more covariant regions of objects in said images and computing a vector in respect of each of said regions based on the appearance of the respective region.

The Applicant goes on to argue that the vectors used in Jain do not include descriptors and vector clustering is not performed in relation to them and then goes on to state that descriptor is a technical term which is a 2^7 dimensional vector which represents an affine invariant region, the implementation of the descriptor being known

in the prior art. The Examiner again respectfully disagrees and has several issues to address regarding this argument. First in response to applicant's argument that the references fail to show the descriptor of the applicant's invention, it is noted that the features upon which applicant relies (i.e., that descriptor is a 2^7 dimensional vector which represents an affine invariant region) are not recited in the rejected claims. Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993). Although indeed the definition given in the specification is one type of descriptor for a vector, it is not the only type and thus the Examiner gives the broadest reasonable interpretation to the term descriptor. If the Applicant wishes the descriptor to be read in accordance with the definition given in the specification then it should be included in the claim language. Furthermore for the sake of argument, Jain discloses that the visual feature is any property of an image that can be computed using computer-vision, and the Applicant admits that the descriptor in the specification is prior art in object retreival, thus even if this recitation were to be included in the claim language it would still be covered by Jain.

The Applicant further argues that the visual senses in Jain relate to the appearance of objects whereas the visual aspects of claim 16 relate to the orientation of an object within frames. Again the Examiner respectfully disagrees. Jain specifically discloses in col. 10 line 22 that the visual feature can relate to orientation, which reads on the limitations of claim 16.

Finally in addressing the 103 rejection the Applicant again relies on features that are not claimed such as the object retrieval which is substantially orientation invariant, and not using semantic level processing. Thus these arguments are irrelevant to the claim language.

Claim Rejections - 35 USC § 102

3.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

4.   Claims 1-5, 7-8, and 11-16 are rejected under 35 U.S.C. 102(b) as being anticipated by Jain et al (US 5,983,237).

Consider claim 1, Jain discloses a method of identifying a user-specified object contained in one or more images of a plurality of images (see column 7, lines 10-23 where Jain describes an image searching method involving a user query), the method comprising defining one or more covariant regions of objects in said images (see column 10, lines 6- 26 describing each image containing visual senses also called visual objects), computing a vector in respect of each of said regions based on the appearance of the respective region, each said vector comprising a descriptor representing that co-variant region (see column 10, lines 16-26 describing a feature vector describing a visual object using covariant descriptors such as texture measures, hue, etc.), vector quantizing said descriptors into clusters (see column 10, lines 43-47 describing grouping the feature vectors to cover a region), storing said clusters as an

index with the images in which they occur (see column 2, line 57 to column 3, line 12

describing insertion of images and associated feature vectors into a database), defining

one or more co-variant regions of said user-specified object (see column 10, lines 24-26

describing computing features of a small region on an image), computing a vector in

respect of each of said regions based on the appearance of said regions, each said

vector comprising a descriptor representing that co-variant region, and vector quantizing

said descriptors into said clusters (see column 11, lines 19-23 describing obtaining

feature vectors for a query image), searching said index and identifying which of said

plurality of images contains said clusters so as to return the images containing said

user-defined object (see column 11, lines 24-31 describing the comparison of feature

vectors and returning a ranking of image matches from a database).

Consider claim 16, Jain discloses a method of identifying a user-specified object

contained in one or more image frames of a moving picture (see column 7, lines 10-23

where Jain describes an image searching method involving a user query, and column 9,

line 64 to column 10, line 3 describing the extension of this search to video), the method

comprising associating a plurality of different "visual aspects" with each of a plurality of

respective objects in said moving picture wherein the visual aspects pertain th the

viewpoint of the object (see column 10, lines 6- 27 describing the process of indexing

images based on the features detected within that image such as orientation), retrieving

the "visual aspects" associated with said user-specified object (see column 11, lines 19-

23 describing submitting the user query and retrieving synonymous feature vectors),

and matching said "visual aspects" associated with said user-specified object with

objects in said frames of said moving picture so as to identify instances of said user-specified object in said frames (see column 11, lines 23-31 describing matching the synonym feature vectors with images in a database and sending back a ranking of the hits).

Consider claim 2, Jain discloses comparing the clusters relating to the objects contained in the images identified as containing an occurrence of said user-specified object with the one or more clusters relating to said user-specified object, and ranking said images identified as containing an occurrence of said user- specified object according to the similarity of the one or more clusters associated therewith to the cluster associated with said user-specified object (see column 11, lines 24-31 describing the comparison of feature vectors and returning a ranking of image matches from a database).

Consider claim 3, Jain discloses that at least two types of viewpoint covariant regions are defined in respect of each of said images (see column 11, lines 19-23 describing submitting the user query and retrieving equivalent query synonyms).

Consider claim 4, Jain discloses that a descriptor is computed in respect of each type of viewpoint covariant region (see column 11, lines 19-23 describing that the equivalent query synonyms are represented by feature vectors).

Consider claim 5, Jain discloses that one or more separate clusters are formed in respect of each type of viewpoint covariant region (see column 10, lines 43-47 describing grouping the feature vectors to cover a region).

Consider claim 7, Jain discloses that said user-specified object is specified as a

sub-part of an image (see column 17, lines 47-53 and figure 6, describing feature

spaces in an image as being disjoint regions).

Consider claim 8, Jain discloses that identification of said user-specified object is

performed by first vector quantizing the descriptor vectors in a sub-part of an image to

precomputed cluster centers (see column 10, lines 43-54 describing grouping the

feature vectors to cover a region or using one large feature to cover the full feature

region).

Consider claim 11, Jain discloses that each image or portion thereof is

represented by one or more cluster frequencies (see column 8, line 66 to column 9, line

3 describing the association of a set of feature vectors to describe the visual

appearance of an image).

Consider claim 12, Jain discloses that said cluster frequency is weighted (see

column 9, lines 3-10 describing weights being associated with the sets of feature

vectors).

Consider claim 13, Jain discloses that a predetermined proportion of most

frequently occurring clusters in said plurality of images are omitted from or suppressed

in such index (see column 15, line 55 to column 16, line 14 describing a diversity

maximization process that limits results by using match quotas).

Consider claim 14, Jain discloses that said index comprises an inverted file

structure having an entry for each cluster which stores all occurrences of the same

cluster in all of said plurality of images and possibly more precomputed information

about each cluster occurrence such as for example its spatial neighbours in an image

(see column 17, line 45 to column 18, line 23 describing indexing feature vector groups

by storing images together that are represented by the same feature vector cluster).

Consider claim 15, Jain discloses including the step of ranking said images using

local image spatial coherence or global relationships of said descriptor vectors (see

column 10, lines 21-25 describing features such as orientation, shape, and turning

angle histograms and column 11, lines 32-43 describing ranking the image hits based

on weightings of each feature vector).

Claim Rejections - 35 USC § 103

5.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

6.   Claims 9-10 and 17-21 are rejected under 35 U.S.C. 103(a) as being

unpatentable over Jain as applied to claims 1 and 16 above, and further in view of

Crabtree et al (US 6,263,088).

Consider claim 9, Jain discloses the method according to claim 1, where images

are grouped using any of the techniques known in Computer Vision and Patter

Recognition research at the time of invention (see column 22, lines 30-35). Jain does

not explicitly disclose that the regions defined in each image are tracked through

contiguous images and unstable regions are rejected. Crabtree discloses a

correspondence graph manager which creates video tracks from regions of motion (see

column 5, lines 20-22) and a region corresponder which outputs a score of the correspondence between regions of a previous frame, when this score is too low it tells the correspondence graph manager to stop tracking the current object (column 24, lines 24-32).

It would have been obvious to one skilled in the art at the time the invention was made to modify the invention of Jain, and modify the region detection to include tracking of images, as taught by Crabtree, thus using a grouping technique known in the art at the time of invention, as discussed by Jain (see column 22, lines 30-35).

Consider claim 10, Crabtree discloses that an estimate of a descriptor for a track is computed from the descriptors throughout the track (see column 23, lines 42-50 describing computing a mean vector from the set of data points whose cluster was matched from one frame to another).

Consider claim 17, Jain discloses the method according to claim 16, wherein the "visual aspects" associated with an object are obtained using any of the techniques known in Computer Vision and Patter Recognition research at the time of invention (see column 22, lines 30-35). Jain does not explicitly disclose using one or more sequences or shots of a moving picture in which said object occurs. Crabtree discloses a method for breaking a video into tracks representing the motion of detected objects (see column 2, line 65 to column 3, line 19).

It would have been obvious to one skilled in the art at the time the invention was made to modify the invention of Jain, and modify the feature detection to operate on video clips of specific objects, as taught by Crabtree, thus using a grouping technique

known in the art at the time of invention, as discussed by Jain (see column 22, lines 30-35).

Consider claim 18, Jain discloses the method according to claim 16. Jain does not explicitly disclose tracking said object through a plurality of image frames in a sequence. Crabtree discloses a method tracking movement of objects through a video scene (see column 2, line 65 to column 3, line 1).

It would have been obvious to one skilled in the art at the time the invention was made to modify the invention of Jain, and modify the feature detection to operate on video clips of specific objects that have been tracked through a video sequence, as taught by Crabtree, thus allowing complex objects to be tracked in an inexpensive manner, as discussed by Crabtree (see column 2, lines 58-63).

Consider claim 19, Crabtree discloses defining affine invariant regions of objects in said image frames and tracking one or more regions through a plurality of image frames in a sequence (see column 18, line 59 to column 19, line 7 describing features used for image tracking, including moment invariant features).

Consider claim 20, Crabtree discloses that in the event that a track terminates in an image frame of a sequence, propagating the track to either following or preceding image frames in the sequence, so as to create a substantially continuous track throughout the image frames in the sequence (see column 5, lines 20-22 describing a correspondence graph manager which creates video tracks from regions of motion and column 24, lines 24-32 describing a region corresponder which outputs a score of the

correspondence between regions of a previous frame, when this score is too low it tells the correspondence graph manager to stop tracking the current object).

Consider claim 21, Crabtree discloses tracked regions being grouped into objects according to their common motion using constraints arising from rigid or semi- rigid object motion (see column 25, lines 57-67 describing a split/merge resolver which uses motion features of regions to calculate confidence values between frames, thus determining whether an object is the same as the previous frame).

Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over Jain as applied to claim 3 above, and further in view of Schaffalitzky et al, "Multi-view matching for unordered image sets" and Matas et al, "Robust Wide Baseline Stereo from Maximally Stable Extremal Regions".

Consider claim 6, Jain discloses the method according to claim 3, where images are grouped using any of the techniques known in Computer Vision and Patter Recognition research at the time of invention (see column 22, lines 30-35). Jain does not explicitly describe using at least two types of viewpoint covariant regions including Shape Adapted and Maximally Stable regions respectively. Schaffalitzky discloses a shape adapted method to extract viewpoint covariant regions (see page 4 describing invariant neighbourhoods). Matas discloses a maximally stable method to extract viewpoint covariant regions (see page 386 describing maximally stable extremal regions).

It would have been obvious to one skilled in the art at the time the invention was made to modify the invention of Jain, and modify the detection of viewpoint invariant

regions to use shape adapte and maximally stable regions, as taught by Schaffalitzky

and Matas, thus using grouping techniques known in the art at the time of invention, as

discussed by Jain (see column 22, lines 30-35).

Conclusion

THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time

policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the mailing date of this final action.

Contact Information

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to JOHN B. STREGE whose telephone number is

(571)272-7457. The examiner can normally be reached on Monday-Friday between the

hours of 8-5.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Matthew Bella can be reached on (571) 272-7778. The fax phone number

for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Matthew C Bella/
Supervisory Patent Examiner, Art
Unit 2624

JS